IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF PENNSYLVANIA

CARNEGIE MELLON UNIVERSITY, Plaintiff, v. MARVELL TECHNOLOGY GROUP, LTD., and MARVELL SEMICONDUCTOR, INC., Defendants.	)	Civil Action No. 09-290 Judge Nora Barry Fischer

OPINION

I. BACKGROUND

This is a patent infringement action brought by Carnegie Mellon University ("CMU") against Marvell Technology Group, LTD., and Marvell Semiconductor, Inc. (collectively "Marvell"). CMU filed this case on March 6, 2009, alleging the infringement of U.S. Patent No. 6,201,839 ("'839 Patent") and U.S. Patent No. 6,438,180 ("'180 Patent"). (Docket No. 1). The technology at issue in this case is generally related to methods and devices for reading information off high density magnetic recording devices, and more specifically, high density hard disk drive sequence detectors. *See* '839 Patent col. 1 ln. 20-23. The disputed claims at issue in this opinion involve the detection of information that is subject to noise and the ways the noise can be accounted for in the sequence detection. *Id.* at col. 2 ln. 1-32.

On January 6, 2010, the parties filed a "Joint Agreed and Disputed Claim Terms Chart" pursuant to LPR 4.2 of the Local Patent Rules for the United States District Court for the Western District of Pennsylvania.[1] (Docket No. 74). Opening claim construction briefs were

---

[1]The Local Patent Rules for the United States District Court for the Western District of Pennsylvania can be found at http://www.pawd.uscourts.gov/Documents/Forms/lrmanual.pdf

filed on January 27, 2010 and February 17, 2010 by CMU and Marvell respectively, along with declarations and exhibits in support. (Docket Nos. 78-84). CMU filed a reply brief and declarations in support on March 3, 2010 and Marvell filed their sur-reply brief and declarations in support on March 11, 2010. (Docket Nos. 89-91, 93-95). A technology tutorial was held on April 7, 2010 and a *Markman* claim construction hearing was held on April 12 and 13, 2010. (Docket Nos. 104-106). Pursuant to an agreement reaching during the claim construction hearing, on May 14, 2010, the parties filed their "Revised Joint Agreed and Disputed Claim Terms Chart" ("RDCTC"). (Docket No. 120). The parties both filed their post-hearing claim construction briefs along with supporting exhibits and declarations on May 28, 2010. (Docket Nos. 128-130). On June 11, 2010, CMU filed a motion for leave to file a post-hearing reply brief, requesting the opportunity to respond to allegedly new extrinsic evidence cited in Marvell's post-hearing brief. (Docket No. 135). Marvell countered with a brief in opposition to CMU's motion for leave and submitted a supporting declaration on June 21, 2010. (Docket Nos. 138, 139). Two days later, the Court granted CMU's motion and ordered that Marvell's brief would be considered as a sur-reply brief to CMU's post-hearing reply brief. Given the complexity of the issues presented by this case and after hearing from the parties, through counsel[2], the Court appointed Dr. Daniel Costello[3] as a technical advisor[4] to assist the Court with

---

(last updated December 1, 2009).

[2] (*See* Docket Nos. 117, 121, 122, 136, 140, 145).

[3] Dr. Costello has a Ph. D. in electrical engineering from the University of Notre Dame and is Professor Emeritus of electrical engineering at the University of Notre Dame. *See* http://xml.ee.nd.edu/faculty/faculty.xsp?id=64636f7374656c31 (last visited 9/27/2010).

[4] The inherent power of the Court allows for the appointment neutral advisors to assist judges in the performance of specific judicial duties. *See Ex parte Peterson*, 253 U.S. 300, 312 (1920); *In re Kensington Int'l. Ltd.*, 368 F.3d 289, 320 (3d Cir. 2004) (Fuentes, J. dissenting); *Techsearch, L.L.C. v. Intel Corp*., 286 F.3d 1360, 1377- 79 (Fed. Cir. 2002); *Eash v. Riggins Trucking Inc.*,

its understanding of the technology on June 28, 2010. (Docket No. 146). The Court held telephone conferences with the technical advisor on July 28, August 11, and August 20, 2010. (Docket Nos. 150, 152, 159). In light of the record as described above, the Court hereby issues the following construction for the disputed claim terms.

II. Legal Standard

"It is a 'bedrock principle' of patent law that 'the claims of a patent define the invention to which the patentee is entitled the right to exclude.'" *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005)(en banc)(citing *Innova/Pure Water, Inc. v. Safari Water Filtration Systems, Inc.*, 381 F.3d 1111, 1115 (Fed. Cir. 2005)). In patent infringement litigation, the court is to first determine, as a matter of law, the proper construction, or meaning, of the disputed claims. Once the claim terms have been properly construed, the fact finder must then determine whether the accused product or method infringes the asserted claims as so construed. *See Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 377-90 (1996).

The decision by the Court of Appeals for the Federal Circuit in *Phillips* provides a blueprint for the court's claim construction analysis:

> A court construing a patent claim seeks to [afford] a claim the [ordinary and customary] meaning it would have to a person of ordinary skill in the art at the time of the invention....
>
> In some cases, the ordinary meaning of claim language as understood by a person of skill in the art may be readily apparent even to lay judges, and claim construction ... involves little more than the application of the widely accepted meaning of commonly understood words ....

---

757 F.2d 557, 563 (3d Cir. 1985). See also *In re Kensington*, 368 F.3d 289, 305 ("[A] judge may consult ex parte with a disinterested expert provided that the judge 'gives notice to the parties of the person consulted and the substance of the advice, and affords the parties a reasonable opportunity to respond.'") (quoting Code of Conduct for U.S. Judges Canon 3 § A(4) (2003)).

> In many cases ..., however, determining the ordinary and customary meaning of the claim requires examination of terms that have a particular meaning in a field of art. Because the meaning ... as understood by persons of skill in the art is often not immediately apparent, and because patentees frequently use terms idiosyncratically, the court looks to those sources available to the public that show what a person of skill in the art would have understood disputed claim language to mean .... Those sources include the words of the claims themselves, the remainder of the specification, the prosecution history, and extrinsic evidence concerning relevant scientific principles, the meaning of technical terms, and the state of the art ....
>
> Within the class of extrinsic evidence, the court has observed that dictionaries and treatises can be useful in claim construction .... We have especially noted the help ... technical dictionaries may provide to a court to better understand the underlying technology and the way in which one of skill in the art might use the claim terms .... Because dictionaries, and especially technical dictionaries, endeavor to collect the accepted meanings of terms used in various fields of science and technology, those resources have been properly recognized as among the many tools that can assist the court in determining the meaning of particular terminology to those of skill in the art .... Such evidence, we have held, may be considered if the court deems it helpful in determining the true meaning of [the] language used . . . .
>
> [Although] ... extrinsic evidence in [a] general [sense is] less reliable than [intrinsic evidence] in determining how to read claim terms, ... encyclopedias and treatises [can be] particularly useful resources to assist the court in determining the ordinary and customary meanings of claim terms [so long as they are considered within the context of the intrinsic evidence].

Phillips, 415 F.3d at 1303, 1313, 1314, 1318-19 (Fed. Cir. 2005) (citations and internal quotations omitted).

In first looking to the intrinsic evidence, the court considers the language of the claim, the specification and the prosecution history. *Vitronics Corp. v. Conceptronics, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996). The claim itself, of course, is of primary importance since "it is that language that the patentee chose to use to 'particularly point[ ] out and distinctly [ ] claim

the subject matter which the patentee regards as his invention.'" *Interactive Gift Express, Inc. v. Compuserve, Inc.*, 231 F.3d 859, 865 (Fed. Cir. 2000) (quoting 35 U.S.C. § 112, ¶ 2).

As to the specification, the United States Court of Appeals for the Federal Circuit has stated that "[it] 'is always highly relevant to the claim construction analysis. Usually it is dispositive; it is the single best guide to the meaning of the disputed claim.'" *Phillips*, 415 F.3d at 1315 (quoting *Vitronics*, 90 F.3d at 1582). It has been warned, however, that there is a danger of reading limitations from the specification into the claim. *Phillips*, 415 F.3d at 1323, (citing *Comark Communications, Inc. v. Harris Corp*., 156 F.3d 1182, 1186 (Fed. Cir.1998) ("there is sometimes a fine line between reading a claim in light of the specification, and reading a limitation into the claim from the specification")). Thus, although the specification may be the best guide in interpreting a disputed term, the court must be careful to use the specification only to ascertain its meaning and not to impose a limit on a claim term. *Id. See also Abbott Laboratories v. Sandoz, Inc.,* 566 F.3d 1282, 1288 (Fed. Cir. 2009) (courts may not limit broad claim language to that described in even a single embodiment absent a clear intention by the patentee to so limit the claim scope); *Karlin Technology, Inc. v. Surgical Dynamics, Inc.,* 177 F.3d 968, 973 (Fed. Cir. 1999) ("The general rule . . . is that the claims of a patent are not limited to the preferred embodiment" of the invention described in the specification).

The final piece of intrinsic evidence is the prosecution or file history. The prosecution history "consists of all express representations made by or on behalf of the applicant to the [patent] examiner to induce a patent grant." *Howmedica Osteonics Corp. v. Wright Medical Technology, Inc.*, 540 F.3d 1337, 1346 (Fed. Cir. 2008) (internal quotations and citations omitted). The prosecution history, therefore, can grant some insight into the scope of the invention, as understood by the inventor and the USPTO, and whether that scope is narrower

than the claim language would otherwise indicate. *Phillips*, 415 F.3d at 1317. A caveat to the prosecution history is that it "often lacks the clarity of the specification and thus is less useful for claim construction purposes." *Id*. As a result, in order for a limitation of the scope of the claims to be read from the prosecution history, the inventor must have made a "clear and unmistakable disavowal" of a broader scope of protection during prosecution. *See Purdue Pharma L.P. v. Endo Pharms. Inc*., 438 F.3d 1123, 1136 (Fed. Cir. 2006).

Lastly, in regards to extrinsic evidence, if the meaning of a claim's terms cannot be ascertained through the intrinsic evidence a court may look to extrinsic evidence, "which 'consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises.'" *Phillips*, 415 F.3d at 1317 (quoting *Markman*, 52 F.3d 967, 980 (Fed. Cir. 1995)). Although extrinsic evidence "can shed useful light on the relevant art," *Phillips*, 415 F.3d at 1317 (quoting *C.R. Bard, Inc. v. U.S. Surgical Corp*., 388 F.3d 858, 862 (Fed. Cir. 2004) (internal quotation omitted)), such evidence is less significant in defining a claim term than intrinsic evidence. This is because "[i]n most situations, an analysis of the intrinsic evidence alone will resolve any ambiguity in a disputed claim term," and "[i]n such circumstances, it is improper to rely on extrinsic evidence." *Vitronics,* 90 F.3d at 1583. Therefore, extrinsic evidence "may be used only to help the court come to the proper understanding of the claims; it may not be used to vary or contradict the claim language" or the other intrinsic evidence. *Id*. at 1584.

III. ANALYSIS

The '839 and '180 Patents are entitled "Method and Apparatus for Correlation-Sensitive Sequence Detection" and "Soft and Hard Sequence Detection in ISI Memory Channels," respectively. '839 Patent col. 1 ln. 1-3; '180 Patent col. 1. ln. 1-2. The '839 Patent was filed on

April 3, 1998 and issued on March 13, 2001. The ‘180 Patent was filed on March 1, 1999 and issued on August 20, 2002. A portion of the ‘180 Patent shares a common specification with the ‘839 Patent. The ‘180 Patent issued from a continuation-in-part application that is related to the application that became the ‘839 Patent. The parties do not dispute that the claim terms used in both applications share a common meaning. For that reason, unless otherwise specified, the following discussion will apply to both the ‘839 and ‘180 Patents.

A. **Correlation**; **Correlation-sensitive branch metrics**; **Correlation-sensitive metric computation update circuit**

The principal dispute between the parties for the “correlation” terms is whether “correlation” refers to a general English meaning of relatedness or if it refers to a specific statistical usage and calculation as found in technical dictionaries.

The term “correlation” as well as the terms “correlation-sensitive branch metrics” and “correlation-sensitive metric computation update circuit” appear in claims 11, 16, 19, and 23 of the ‘839 Patent and claim 6 of the ‘180 Patent. Claim 11 of the ‘839 Patent is illustrative of how the terms are used and it reads as follows:

> **11.** A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols stored on a high density magnetic recording device, comprising the steps of:
> (a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;
> (b) outputting a delayed decision on the recorded symbol;
> (c) outputting a delayed signal sample;
> (d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and delayed decisions;
> (e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and
> (f) repeating steps (a)-(e) for every new signal sample.

‘839 Patent col. 15 ln. 2-17.

CMU and Marvell agree on the construction of “correlation-sensitive metric computation update circuit” except for the underlying definition of “correlation” and “correlation sensitive branch metrics.” The parties’ agreed construction is:

> “A correlation sensitive metric computation update circuit” means “a circuit that recalculates ‘correlation*-sensitive branch metrics’** using statistics from the ‘noise statistics tracker circuit.’”

RDCTC at 8.

CMU’s constructions of “correlation” and “correlation-sensitive branch metrics” are:

> “Correlation” means “the degree to which two or more items (here, noise in signal samples) show a tendency to vary together.”
>
> “Correlation sensitive branch metrics” means “‘branch metrics’ that use ‘correlation’ in signal samples in their calculation by including at least one term that involves multiplying signal samples from different time instances.”

RDCTC at 6-7.

Marvell’s construction of “correlation” and “correlation-sensitive branch metrics” are:

> “Correlation” means “the expected (mean) value of the product of two random variables (e.g. $E[r_i r_j]$, where r_i and r_j are signal samples at time i and j, respectively).”
>
> “Correlation sensitive branch metrics” means “‘branch metrics’ that use ‘correlation’ in signal samples in their calculation.”

RDCTC at 6-7.

As an initial matter, Marvell argues that its definition of “correlation” ought to be adopted because it reflects how the person having ordinary skill in the art (“PHOSITA”) would understand the use of the term based in part on the fact that its definition comes from a technical dictionary and thus is closest to the term’s ordinary meaning in engineering and statistics.

Although CMU cited several general English dictionaries[5] in support of its construction, CMU also cited to three statistics dictionaries. The definitions from the statistics dictionaries cited by CMU are as follows:

> **Correlation:** A general term for interdependence between pairs of variables.
>
> **Correlation** A general term used to describe the fact that two (or more) variables are related . . . Although the word 'correlation' is used loosely to describe the existence of some general relationship, it has a more specific meaning in the context of linear relations between variables. *See* correlation coefficient.
>
> **Correlation**- A general term denoting association or relationship between two or more variables. More generally, it is the extent or degree to which two or more quantities are associated or related. It is measured by an index called correlation coefficient.

Cambridge Dictionary of Statistics, 3rd ed. (2006); *Oxford Dictionary of Statistics* (2002); *Pocket Dictionary of Statistics*[6] (2002). (Docket No. 79-9, p. 9-20).

Marvell, in support of its construction, cites to several statistics textbooks that state the following about "correlation:"

> In electrical engineering, it is customary to call the $j = 1$ $k = 1$ moment, E[XY], the **correlation of X and Y**.
>
> The second-order moment m_{11} = E[XY] is called the *correlation* of X and Y. It is so important to later work that we give it the symbol R_{xy}.

---

[5] The general English dictionaries that CMU cited are the following: *Merriam-Webster's Collegiate Dictionary*, 10th ed (1996); *The Random House Dictionary of the English Language*, (1996); *Webster's New Universal Unabridged Dictionary*, 2nd ed. (1983); *The Compact Oxford English Dictionary*, 2nd ed. (1987); and *Webster's Third New International Dictionary* (1996). (Docket No. 79-1, 79-2).

[6] This dictionary was cited by Marvell for the construction of "noise covariance matrices."

> The probabilistic *autocorrelation of the random process* X, at the times t_1 and t_2, is simply the correlation of the two random variables $X(t_1)$ and $X(t_2)$, and is denoted by
> $$E\{X(t_1)X(t_2)\} = R_x(t_1,t_2).$$
>
> **Correlation** describes a random process in a way that is impossible for the mean and the variance to do. Both the mean and variance depend on a pdf, which included time at only one instant. Correlation, on the other hand, is a bivariate parameter, and uses a bivariate pdf [] which included two different times: t_1 and t_2. We therefore define an *autocorrelation function*:
> $$R_x(t_1,t_2) = E[X(t_1)X(t_2)]$$

Leon-Garcia, *Probability and Random Processes for Electrical Engineering* (2nd ed. 1994) (Docket No. 82-18); P. Peebles, *Probability, Random Variables, and Random Signal Principles*, (1980) (Docket No. 83-2); W. Gardner, *Introduction to Random Processes with Applications to Signal and Systems*, (1986) (Docket No. 83-3); R.H. Williams, *Electrical Engineering Probability*, (1991) (Docket No. 83-4).

CMU and Marvell agree that the PHOSITA in this case would be a person with at least a Master's degree in electrical engineering who had specialized in data detection and signal processing and had at least two years work experience in the industry. *See* (Docket No. 79-1, Dr. McLaughlin Declaration at ¶8); (Docket No. 84, Dr. Proakis Declaration at ¶29). Based upon the above definitions and the agreed upon qualifications of the PHOSITA, the Court concludes that the PHOSITA would be knowledgeable about both the general definition of "correlation" used by CMU as well as the more technical definition and formula used by Marvell. The question then is not which definition would the PHOSITA understand and use in the daily practice of his or her art. Rather, the issue is which definition would the PHOSITA conclude to be the most consistent with the claims, specification and prosecution history (the intrinsic evidence) of the '839 and '180 Patents. The above cited extrinsic evidence is thus only relied on

to the extent that it informs the Court as to what the PHOSITA's general understanding is; however, the following treatment of the claim terms is based on only the intrinsic record.

As stated in the preamble of claim 11 of the '839 Patent, the method "exploits the correlation between adjacent signal samples." '839 Patent col. 15 ln. 2-3. The purpose for exploiting the correlation is explained in the background and summary of the invention section of the '839 Patent. In the background of the invention section of the specification, it is stated that the "[prior art] methods do not take in consideration the correlation between noise samples in the readback signal. These correlations arise due to noise coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities." '839 Patent col 1 ln. 57-61. In the summary of the invention section of the '839 Patent, it states:

> In high density magnetic recording, noise samples corresponding to adjacent signal samples are heavily correlated as a result of front-end equalizers, media noise, and signal nonlinearities combined with nonlinear filters to cancel them. This correlation deteriorates significantly the performance of detectors at high densities.
>
> The trellis/tree branch metric computation of the present invention is correlation-sensitive, being both signal dependent and sensitive to correlations between noise samples.

'839 Patent col. 2 ln. 2-11.

From this, a PHOSITA would conclude that there is a correlation between noise samples, that the correlation exists due to certain structures that exist in the recording circuit, and that the failure to account for the correlation between the noise samples in the prior art detracts from the performance of the prior art detectors. How the correlation between the noise samples is taken into account and "exploited" is what is discussed in the remainder of the specification.

The detailed description of the invention section of the '839 Patent discusses the prior art branch metrics before explaining how the "Correlation-sensitive branch metric" works. For the

prior art Euclidian branch metric, it is stated, that “the noise samples are realizations of *independent* identically distributed Gaussian random variables with zero mean and variance σ^2... This implies that the correlation distance is L = 0 and that the noise pdfs have the same form for all noise samples.” ‘839 Patent col. 5 ln. 60-64 (emphasis added). L is stated to be the correlation length of the noise, that is, the number of signal samples with which the noise is considered to be correlated. (*See* Docket No. 79-1, McLaughlin Declaration at ¶11). As a result, the Euclidian branch metric does not consider any of the samples to be correlated with each other, therefore, they are independent and the correlation length or distance is set at zero.

For the prior art variance dependent branch metric, it is stated that “the noise samples are samples of *independent* Gaussian variables, but that their variance depends on the written sequence of symbols. The noise correlation length is still L = 0, but the variance of the noise samples is no longer constant for all samples.” ‘839 Patent col. 6 ln. 16-20 (emphasis added). Since the variance dependent branch metric accounts for noise that depends on the written sequence of symbols, this branch metric is said to be signal-dependent. The variance dependent branch metic, however, still has its correlation length set to zero, thus it does not take into account noise from other signal samples.

The correlation-sensitive branch metric is first distinguished from the prior art Euclidian and variance-dependent branch metrics by indicating that the correlation length is now set at some number greater than zero. ‘839 Patent col. 6 ln. 36-37. As a result, the correlation-sensitive branch metric uses more than one signal sample from different time instances so as to take into account the noise at a given time that is attributable to noise from other time instances. (*See* Docket No. 79-1, McLaughlin Declaration at ¶11). The specification goes on to explain

how the correlation-sensitive branch metric can be calculated and it is provided that the general correlation-sensitive branch metric is equation (13). '839 Patent col 6 ln. 66 - col. 7 ln. 5.

As CMU points out, in one of the embodiments, the general correlation sensitive branch metric (13) is calculated, but the correlation equation cited by Marvell is not used. According to CMU, the variance (σ^2), which is used in the first logarithmic term of equation (13), is shown to be able to be calculated without using the equation cited by Marvell for correlation. *See* (Docket No. 90-3, McLaughlin 2nd Declaration at ¶9). CMU points to equation (23) to show that the variance used in logarithmic circuit **50** may be calculated without a correlation value or the use of a covariance matrix. '839 Patent col. 7 ln. 14-15; col. 10 ln. 12-13. Marvell does not dispute this, but rather argues that a claim need not cover every embodiment. (Docket No. 93 at 6-7)(citing *Intamin*, 483 F.3d at 1337). However, it is the Court's conclusion that the PHOSITA would not read the general correlation-sensitive branch metric equation (13) to become correlation insensitive if equation (23) were utilized in calculating the branch metric. Equation (13) would still account for the noise at a given time that is attributable to noise at other time instances with the utilization of equation (23). As a result, since Marvell's construction would require the use of calculating the expected (mean) value of the product of two random variables in the correlation-sensitive branch metric, it is not supported by the specification.

Support for CMU's general English construction can also be found in the detailed description. For example, a PHOSITA would read the following language to refer to the general meaning of correlation and not the formula provided by Marvell:

> Although recently very popular, such a method [of employing a nonlinear filter] introduces further correlation between noise samples due to the nonlinear character of the filter.
>
> Note that both filters add correlation to the noise.

> The PR4(C2) detector performs better because it partially removes the effects of noise correlation introduced by the PR4 shaping filter.
>
> [T]he EPR4 shaping filter does not introduce unnecessary noise correlation.
>
> PR4(C2) still outperforms the two other algorithms, showing the value of exploiting the correlation across signal samples.

'839 Patent col. 9 ln. 8-11; col 12 ln. 35; col 12 ln. 55-57; col 12 ln. 64-65; col 13 ln. 5-7. No mathematical formula is used in reference to these statements and inserting the formula used in Marvell's construction would not be consistent with the meaning of these statements. For example, a PHOSITA would not read that the filters add correlation to the noise to mean that a value from the correlation formula is mathematically added to some value of noise by the filter. Rather, the sentence means that the filter's implementation physically causes noise in the signal sample to vary together to a greater degree.

The prosecution history adds further clarification as to what makes the branch metric in the patent correlation-sensitive. In the prosecution history, the examiner initially rejected claims as being anticipated by U.S. Patent No. 5,862,192 to Huszar et al. (Docket No. 83-1 at 9). The patentee stated that Huszar did not have a correlation-sensitive branch metric because, although the branch metric did contain signal samples from different time instances, the branch metric did not have a term that "corresponds to the correlation between $r_i(0)$and $r_i(1)$, i.e. there is no term that involves multiplying $r_i(0)$ with $r_i(1)$." (Docket No. 83-1 at 10). As a result, the patentee distinguished its branch metric as being correlation-sensitive from Huszar by stating that the signal samples were multiplied together in the branch metric. *Id.*

Marvell takes this statement to mean that the patentee required the calculating of the expected (mean) value of the product of two random variables in its correlation-sensitive branch

metric. CMU points out, however, that there is nothing said in regards to the calculating of the expected value, and that it would be inappropriate to read in the requirement of calculating the expected value of the product of the signal samples. The Court agrees that adding such a limitation would be inconsistent with the specification and how the applicants sought to distinguish the correlation-sensitive branch metric from the branch metric in Huszar. Marvell cites to other sections of the specification of the '839 Patent where the term "expected value" is used; however, those sections do not refer to the expected values of the product of random variables.

In considering all of the intrinsic evidence and extrinsic sources provided by the parties, the PHOSITA would find that "correlation," as used throughout the patent refers to the general English meaning. Marvell argues that CMU's construction comes from an amalgam of sources. The Court, however, is satisfied that CMU's construction both reflects how the PHOSITA would understand the term and is construed in a way to aid a jury in understanding the claims.

While CMU's construction for "correlation sensitive branch metrics" does contain the requirement that multiple signal samples are multiplied together, further clarification should be added to the construction to make the requirement of using more than one signal sample explicit. This requirement would be understood by the PHOSITA as the result of setting the correlation length to a number greater than zero ($L > 0$). To that end, the Court's construction of "correlation sensitive branch metrics" amends CMU's construction to reflect this requirement.

As a result, the Court concludes that the construction of the terms "correlation," "correlation sensitive branch metrics," and "correlation sensitive metric computation update circuit" are as follows:

Correlation means the "degree to which two or more items (here, noise in signal samples) show a tendency to vary together."

Correlation sensitive branch metrics means "'branch metrics' that account for 'correlation' in the signal samples by using multiple signal samples from different time instances and including at least one term in the branch metric calculation that involves multiplying signal samples from different time instances together."

A correlation sensitive metric computation update circuit means "a circuit that recalculates 'correlation-sensitive branch metrics' using statistics from the 'noise statistics tracker circuit.'"

B. Correlated Noise

The principal dispute between the parties for the construction of "correlated" and "correlated noise" is similar to the dispute involving "correlation." Would the term be understood by the PHOSITA to be defined by its general English definition or by its general meaning as used in engineering and statistics dictionaries?

"Correlated" and "correlated noise" appear in claims 2 and 5 of the '839 Patent and claim 1 of the '180 Patent. Those claims read as follows:

> **2.** The method of claim **1** further comprising the step of receiving said signal samples, said signal samples having signal-dependent noise, correlated noise, or both signal-dependent and correlated noise associated therewith.
>
> **5.** The method of claim **4** further comprising the step of receiving said signal samples, said signal samples having signal-dependent noise, correlated noise, or both signal-dependent noise and correlated noise associated therewith.
>
> **1**. A method of determining branch metric values in a detector, comprising:
>
> receiving a plurality of time variant signal samples, the

> signal samples having one of signal-dependent noise, correlated noise, and both signal dependent and correlated noise associated therewith;
>
> selecting a branch metric function at a certain time index; and
>
> applying the selected function to the signal samples to determine the metric values.

'839 Patent col. 14 ln. 3-6, 20-23; '180 Patent col. 15 ln. 39-48.

CMU's proposed construction of "correlated" and "correlated noise" are:

> Two items are "correlated" when they have a tendency to vary together.
>
> "Correlated Noise" means "noise with 'correlation' among 'signal samples,' such as that caused by coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities."

RDCTC at 5.

Marvell's proposed construction of "correlated noise" is:

> "Correlated noise" means "noise having nonzero 'covariance' (see construction of 'covariance' below)."
>
> "Covariance" means "the expected (mean) value of the product of (r_i-m_i) and (r_j-m_j), where r_i and r_j are observed signal samples (at time i and j, respectively) and m_i and m_j are the expected (mean) values of the samples (at time i and j, respectively)(i.e., $E[(r_i-m_i)(r_j-m_j)]$)."

RDCTC at 5, 9.

CMU's sources for its construction of "correlated" are the same dictionaries used to interpret "correlation" as discussed above. *See supra* at 9-10. CMU's source for "correlated noise" is from both its construction of "correlated" and language found in the specifications of the '839 and '180 Patents. Marvell, on the other hand, cites to Dr. Proakis's textbook as the source of its construction for "correlated noise" and "variance." The textbook citation reads as follows:

> Two random variables are said to be *uncorrelated* if $E(X_i X_j) = E(X_i)E(X_j) = m_i m_j$. In that case, the covariance $\mu_{ij} = 0$. We note that when X_i and X_j are statistically independent, they are also uncorrelated. However, if X_i and X_j are uncorrelated, they are not necessarily statistically independent.

Proakis, *Digital Communications* (3d Ed. 1995)(Docket No. 83-14 at 17).

In first looking at the claims, it would be apparent to the PHOSITA that the claim terms distinguish "correlated noise" from signal-dependent noise, but do not otherwise define the term. The term "correlated" appears only twice in the '839 Patent's specification. First, in the summary of the invention section it is stated that "noise samples corresponding to adjacent signal samples are heavily correlated as a result of front-end equalizers, media noise, and signal nonlinearities combined with nonlinear filters to cancel them." '839 Patent col. 2 ln. 3-7. Second, in the detailed description section of the '839 Patent it states that for the correlation-sensitive branch metric "[t]he noise is now considered to be both correlated and signal-dependent." '839 Patent col. 6 ln. 38-39.

CMU points out, and the Court agrees, that the language "heavily correlated" would not make sense if Marvell's construction were adopted as such language could only be understood as meaning that the noise had a heavier or greater nonzero covariance. Such a construction would not be how the PHOSITA would read "heavily correlated." CMU's construction, however, that "two items are 'correlated' when they have a tendency to vary together" would be consistent with the notion that the noise could have a greater or heavier tendency to vary together.

Furthermore, based in part on the construction adopted for "correlation," CMU points out that Marvell's construction is consistent in scope with CMU's, except that it relies on mathematical terminology. The parties agree that if variables are not related then their

covariance is zero.[7] (Docket No. 81 at 7). To say that two random variables have a non-zero covariance is then to say that they have some relation between them or in the words of CMU the "variables will vary together." As a result, stripped of its mathematical terminology, Marvell's construction becomes "correlated noise" means noise that varies together with other noise samples.

Marvell argues that the use of the term "tendency" is too vague. CMU correctly points out, however, that terms of degree, such as "tendency" can be used in claim construction. *Power-One, Inc. v. Artesyn Tech., Inc.*, 599 F.3d 1343, 1348 (Fed. Cir. 2010). Additionally, Marvell's construction is equally as broad, as any non-zero value of the covariance would be construed to be correlated. Therefore, the Court finds no issue with CMU's use of "tendency" in the proposed construction.

Marvell also finds fault with CMU's use of examples of sources of correlated noise as an attempt by CMU to erroneously import limitations into the claims. CMU's construction, however, only uses the sources as examples and does not purport to give a complete list of sources of correlated noise due to use of the words "such as." (*See* Docket No. 119 at 150:18-151:2) Furthermore, the examples are taken directly from the specification and would provide context for the jury to learn what correlated noise is with respect to the Patents.

---

[7] It should be noted that if two random variables have a zero covariance, this does not necessarily mean they are independent variables. "[F]or non-Gaussian random variables: [i]t is possible for X and Y to be uncorrelated but not independent." Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, at 234 (2d ed. 1994)(Docket No. 82-18 at 6). "The correlation of random variables X_1 and X_2 indicates the degree of linear dependence between the variables. If $\rho(X_1, X_2) = 0$, then there is no linear relation between the random variables, but there may well be some different relation between them." Polyanin and Manzhirov, *Handbook of Mathematics for Engineers and Scientists*, at 1061 (2007 ed.) (Docket No. 83-15 at 5).

As a result, the construction for "correlated" and "correlated noise" that is consistent with the intrinsic evidence and would be understood by the PHOSITA is as follows:

Correlated means two items that have a tendency to vary together.

Correlated noise means "noise with 'correlation' among 'signal samples,' such as that caused by coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities."

C. Noise Covariance Matrices

The dispute between the parties for "noise covariance matrices" centers around whether the patentee acted as a lexicographer[8] and gave a definition of "noise covariance matrices" in the specification.

The term "noise covariance matrices" is used in claims 11, 16, 19, and 23 of the '839 Patent and claim 6 of the '180 Patent. Claim 19 of the '839 Patent is illustrative of how the term is used and it reads as follows:

> **19**. A detector circuit for detecting a plurality of data from a plurality of signal samples read from a recording medium comprising:
> a Viterbi-like detector circuit, said Viterbi-like detector circuit for producing a plurality of delayed decisions and a plurality of delayed signal samples from a plurality of signal samples;
> a noise statistics tracker circuit responsive to said Viterbi-like detector circuit for updating a plurality of noise covariance matrices in response to said delayed decisions and said delayed signal samples; and
> a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for

---

[8]"An applicant is entitled to be his or her own lexicographer and may rebut the presumption that claim terms are to be given their ordinary and customary meaning by clearly setting forth a definition of the term that is different from its ordinary and customary meaning(s)." Manual for Patent Examining Procedure §2111.01 IV. (8th ed., rev. July 2010).

recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

'839 Patent col. 15 ln. 50-66.

CMU's construction for "noise covariance matrices" is:

> "Noise covariance matrices" means "noise statistics used to calculate the 'correlation-sensitive branch metrics.'"

RDCTC at 11.

Marvell separately provides constructions of "covariance" and "covariance matrices" in addition to its construction of "noise covariance matrices" all of which are as follows:

> "Covariance" means "the expected (mean) value of the product of (r_i-m_i) and (r_j-m_j), where r_i and r_j are observed signal samples (at time i and j, respectively) and m_i and m_j are the expected (mean) values of the samples (at time i and j, respectively)(i.e., $E[(r_i-m_i)(r_j-m_j)]$)."
>
> "Covariance matrices" means "arrays of covariance of pairs of signal samples, e.g. :
>
> $$\begin{bmatrix} \mathrm{cov}(r_i, r_i) & \mathrm{cov}(r_i, r_{i+1}) \\ \mathrm{cov}(r_{i+1}, r_i) & \mathrm{cov}(r_{i+1}, r_{i+1}) \end{bmatrix}$$
>
> "Noise covariance matrices" means "covariance matrices of signal samples (where the signal samples include noise)."

RDCTC at 9-11.

CMU's source for its construction is from the specification, which CMU argues is an explicit definition for "noise covariance matrices," that reads:

> A noise statistics tracker circuit **34** uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices.

‘839 Patent col. 3 ln. 36-38. CMU then argues that the use of “i.e.” equates “noise covariance matrices” to noise statistics and cites *Abbott Labs v. Novapharm Ltd.*, 323 F.3d 1324 (Fed. Cir. 2003), in support.

Marvell states that its construction is the plain meaning of the term “noise covariance matrices,” and Marvell’s sources for its construction are technical dictionaries and textbooks. Marvell cites the following as an example of the ordinary meaning for “covariance matrix”:

> **covariance matrix** - A square array that represents all the pairs of covariances of a set of random variables. A covariance matrix is a square matrix in which main diagonal elements represent variances of the variables and off-diagonal elements are the covariance. Moreover, like the correlation matrix, a covariance matrix is also symmetrical about the diagonal.

Pocket Dictionary of Statistics at 66 (2002)(Docket No. 82-16 at 4-5).

As stated above, “the specification is ‘the single best guide to the meaning of a disputed term,’” and consistent with that general principle “the specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor’s lexicography governs.” *Phillips*, 415 F.3d 1316, 1321 (quoting *Vitronics*, 90 F.3d at 1582; citing *CCS Fitness, Inc. v. Brunswick Corp.*, 288 F.3d 1359, 1366 (Fed.Cir.2002)). The Court of Appeals for the Federal Circuit has recognized that a patentee can act as a lexicographer with the use of “i.e.[9]” followed by a definition of the term. *Pfizer, Inc. v. Teva Pharmaceuticals USA, Inc.*, 429 F.3d 1364, 1373 (Fed. Cir. 2005). The Court of Appeals cautioned, however, that a person of ordinary skill in the art is always considered to have read the claims in light of the full specification. *Id*. (citing *SanDisk Corp v. Memorex Prods., Inc.*, 415 F.3d 1278, 1285 (Fed. Cir. 2005); *Budde v. Harley-Davidson, Inc.*,

---

[9] “i.e. is an abbreviation for the Latin *id est* and means ‘that is.’” *Webster’s Dictionary of English Usage* (1989)(Docket No. 94-1 at 4).

250 F.3d 1369, 1379-80 (Fed. Cir. 2001)). As a result, support for an alternative construction found elsewhere in the specification may indicate that the patentee did not use "i.e." in a lexicographical manner and thus, the proper construction will be found elsewhere in the patent. *Pfizer*, 429 F.3d at 1373-74.

In *Pfizer,* the Court of Appeals found that "saccharides (i.e. sugars)" did not constitute an explicit definition of saccharides because elsewhere in the specification the patentee explained what saccharides meant. *Id*. at 1373-75. In the specification, there was a section entitled "SACCHARIDES" that gave context to the term that was inconsistent with equating saccharides with sugars only. *Id*. The Court of Appeals stated, however, "[p]roperly understood, then, these sections do not define the exact meaning of 'saccharides' and 'excipients.'" *Id*. at 1374. Rather, since the section left an open ended meaning that was inconsistent with equating saccharides to sugars, the Court of Appeals was left to conclude that the "i.e." was not definitional. As a result, if a person of ordinary skill in the art would find the purported "i.e." definition inconsistent with the rest of the patent, the court will not find the "i.e." to be used to define the term that precedes it. This is consistent with the cases that have found "i.e." to be used to create a definition, because in those cases, either the specification added further support to the definition, or the only use of the term in the specification was in conjunction with the "i.e." *See Abbott*, 323 F.3d at 1330 ("Moreover, the inclusion of the word 'intimate' in the definition, together with the fact that fenofibrate and SLS are the only ingredients present in every co-micronized mixture described in the '726 patent's specification, makes it abundantly clear that [what followed the i.e. is definitional]"); *Tidel Engineering L.P. v. Fire Kind International, Inc.*, 613 F.Supp.2d 823, 829 (E.D. Texas Jan. 6, 2009)("Indeed, the term 'economy safe' is not used but one time in the specification. Following that only instance is the parenthetical '(i.e., comprised of just a safe and

a unit 16, without a PC board and printer).'"); *Caritas Technologies, Inc. v. Comcast Corp.*, No. 2:05-CV-339-DF, 2006 WL 6112191 at *16 (E.D. Texas Oct. 18, 2006)("Therefore, the specification does not teach any other meaning of 'connection status information' besides that taught after 'i.e.' introduces a definition."); *ESN, LLC v. Cisco Sys., Inc.*, No. 5:08-CV-20-DF, 2009 WL 2849742 ("[T]he specification does not identify "SIP User Agent" other than in connection with an endpoint, so the use of "i.e." cited above is especially probative of the meaning of "SIP User Agent.")

In the '839 Patent, the term "noise covariance matrices" appears elsewhere in the patent, not only with the purported "i.e" definition, and it is used as follows:

> Because the noise statistics are non-stationary, the noise sensitive branch metrics are adaptively computed by estimating the noise covariance matrices from the read-back data.
>
> A noise statistics tracker circuit **34** uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices.
>
> The focus is shifted to tracking the noise covariance matrices needed in the computation of the branch metrics (13).

'839 Patent col. 2 ln. 15-18; col. 3 ln. 36-38; col. 9 ln. 21-23.

The term "noise statistics" which CMU argues is equated with "noise covariances matrices" is used elsewhere in the specification, as follows:

> Specific expressions for the branch metrics that result under different assumptions on the noise statistics are next considered.
>
> Also, the signal and noise statistics will be different if a head is flying slightly off-track or if it is flying directly over the track.
>
> The past samples and detector decisions are used to update the noise statistics at step **44**.

'839 Patent col. 5 ln. 56-58; col. 8 ln. 31-33; col. 11 ln. 16-18

In reading the entire patent, the PHOSITA would not read the "i.e." cited by CMU to define "noise covariance matrices" as meaning "noise statistics" because the term's usage elsewhere in the specification indicates they are separate concepts with separate definitions. For instance, the phrase "[t]he focus is shifted to tracking the noise covariance matrices needed in the computation of the branch metrics (13)" would indicate that the noise covariance matrices can be found in equation (13). '839 Patent col. 9 ln. 21-23. The general correlation-sensitive metric equation (13) reads:

$$M_i = \log det \frac{C_i}{det\ \ c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i$$

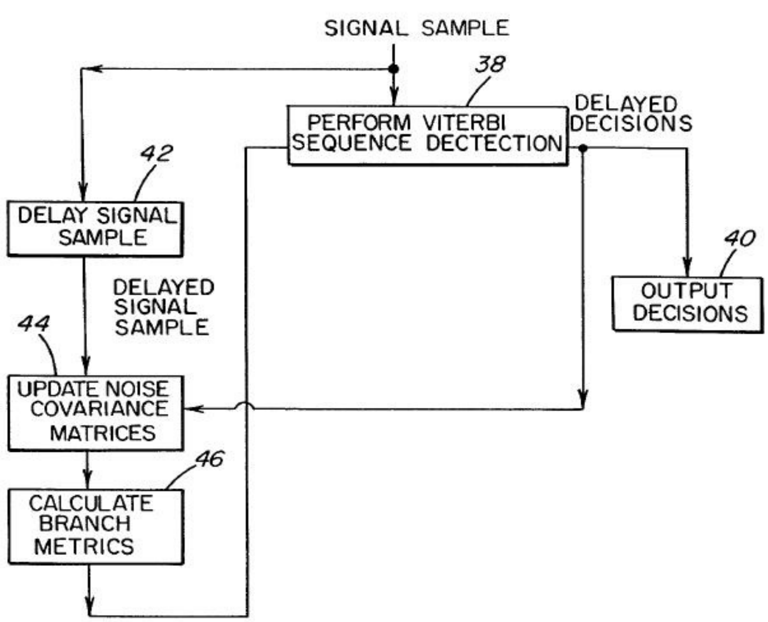
'839 Patent col. 7 ln. 1-5. Prior to equation (13), it is stated that "[t]he (L+1)x(L+1) matrix C_i is the covariance matrix of the data samples r_i, r_{i+1}, . . . , r_{i+L}, when a sequence of symbols a_{i-Kl}, . . . a_{i+L+Kt} is written." '839 Patent col. 6 ln. 53-55. The PHOSITA would take this to mean that the covariance matrix of the data samples C_i used in equation (13) is the "noise covariance matrices" being referred to by the language "the noise covariance matrices needed in the computation of the branch metrics (13)". '839 Patent col. 9 ln. 21-23. Furthermore, the covariance matrices and the "noise covariance matrices" are equated in the summary of the invention section. It states that "[b]ecause the noise statistics are non-stationary, the noise sensitive branch metrics are adaptively computed by estimating the noise covariance matrices from the read-back data. These covariance matrices are different for each branch of the tree/trellis due to the signal-dependent structure of the media noise." '839 Patent col. 2 ln. 15-20.

"Noise statistics" can be shown not to have the same meaning as "noise covariances matrices" in the wording "specific expressions for the branch metrics that result under different assumptions on the noise statistics are next considered." '839 Patent col. 5 ln. 56-58. In this

instance, the assumptions on the "noise statistics" are the prior art assumptions of white Gaussian noise and variance dependent noise as well as the novel assumption that the noise is correlated and signal-dependent for the correlation-sensitive branch metric. Because the term "noise statistics" has a broad enough meaning, as used by the patentee, to encompass both the prior art noise assumptions as well as the correlated and signal-dependent noise assumptions, the intrinsic evidence does not support equating "noise statistics" with "noise covariance matrices" and the PHOSITA would not define "noise covariance matrices" as "noise statistics." To do so would allow the prior art noise assumptions to be used in the correlation-sensitive branch metric which the PHOSITA would understand to be inapposite to the purpose of the invention.

CMU argues that the language "[t]he past samples and detector decisions are used to update the noise statistics at step **44**" provides a second intrinsic example of where the patentee explicitly defines "noise covariances matrices" as "noise statistics." '839 Patent col. 11. ln. 16-18. Step **44** of Fig. 6 shows:

FIG. 6



‘839 Patent Fig. 6. CMU argues that updating the “noise statistics” at step **44** means that “noise statistics” and “noise covariance matrices” are the same thing. The text “[t]he past samples and detector decisions are used to update the noise statistics at step **44**” and Fig. 6 are consistent with the language “[a] noise statistics tracker circuit **34** uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices.” ‘839 Patent col. 11 ln. 16-18; col. 3 ln. 36-38. However, the PHOSITA would not read either section of the Patent to be providing a definition of “noise covariance matrices.” Rather, both statements would be read as meaning when the “noise statistics” are updated, this has the effect of updating the “noise covariance matrices.” This understanding is consistent with the notion that the “noise covariance matrices” are derived from the data samples that contain noise. *See* ‘839 Patent col. 6 ln. 53-55.

CMU also points out that alternative embodiments contained in the specification do not require the use of a covariance matrix in order to calculate the correlation-sensitive branch metric. (*See* Docket No. 90-3, McLaughlin’s 2nd Declaration at ¶¶ 9-12). It is true that “[a] claim construction that excludes a preferred embodiment . . . is ‘rarely, if ever correct.’” *Pfizer*, 429 F.3d at 1374 (citing *SanDisk Corp.*, 415 F.3d at 1285; *Vitronics*, 90 F.3d at 1583). However, at the hearing, Marvell stated that the alternative embodiment mentioned by CMU would be covered by another set of claims, citing specifically claim 20, to which CMU agreed that all embodiments would be covered by at least one claim under Marvell’s construction. (Docket No. 118 at 180:7-12, 204:15-17). “A patentee may draft different claims to cover different embodiments.” *Intamin*, 483 F.3d at 1337. Since the alternative embodiment that calculates the correlation-sensitive branch metric is covered by another claim, the PHOSITA would not have reason to read “noise covariance matrices” as meaning “noise statistics.”

Turning to Marvell's construction, when the Court asked CMU if separate definitions were adopted for "covariance" and "covariance matrices," what would its constructions be, CMU responded that Marvell's constructions were appropriate. (Docket No. 118 at 178:19-179:12). As stated above, the specification says that "C_i is the covariance matrix of the data samples" and C_i is the noise covariance matrix used in equation (13). '839 Patent col. 6 ln. 53-55. This language from the specification is directly reflected in Marvell's construction.

As a result, the construction of the terms "covariance," "covariance matrices and "noise covariance matrices"is as follows:

Covariance means "the expected (mean) value of the product of (r_i-m_i) and (r_j-m_j), where r_i and r_j are observed signal samples (at time i and j, respectively) and m_i and m_j are the expected (mean) values of the samples (at time i and j, respectively)(i.e., $E[(r_i-m_i)(r_j-m_j)]$)."

Covariance matrices means "arrays of covariance of pairs of signal samples, e.g. :

$$\begin{bmatrix} \mathrm{cov}(r_i, r_i) & \mathrm{cov}(r_i, r_{i+1}) \\ \mathrm{cov}(r_{i+1}, r_i) & \mathrm{cov}(r_{i+1}, r_{i+1}) \end{bmatrix}$$

Noise covariance matrices means "covariance matrices of signal samples (where the signal samples include noise)."

D. Signal-Dependent noise; Signal-Dependent branch metric function

Following the claim construction hearing, it is apparent that the parties' dispute over the "signal-dependent noise" terms centers around whether the "signal-dependent noise" must come from the "media noise." *See* (Docket Nos. 128 at 8-10; 129 at 10-11). Both parties agree that "media noise" as used in the patents is limited to magnetic recording. (Docket No. 119 at 25:11-21; 34:4-21; 38:20-23). Furthermore, based on the language, "[t]he non-stationarity of the media noise results from its signal dependent nature" '839 Patent col. 1 ln. 39-41, it would be clear to

the PHOSITA that "media noise" is a type of "signal-dependent noise. The dispute becomes whether "media noise" is the only type of "signal-dependent noise" being referred to in the claims.

"Signal-dependent noise" is used in claims 2 and 5 of the '839 Patent and claim 1 of the '180 Patent. "Signal-dependent branch metric function" is found in claims 3 and 4 of the '839 Patent and claim 2 of the '180 Patent. Claims 2 and 5 of the '839 Patent and claim 1 of the '180 Patent are reproduced above. *Supra* at 16-17. Claim 3 of the '839 Patent is illustrative of how "signal-dependent branch metric function" is used and is as follows:

> **3**. The method of claim **1** wherein said branch metric functions for each of the branches are selected from a set of signal-dependent branch metric functions.

'839 Patent col. 14 ln. 7-9.

CMU's construction of "signal-dependent noise" and "signal-dependent branch metric function" are as follows:

> "Signal-dependent noise" means "media noise in the readback signal whose noise structure is attributable to a specific sequence of symbols (e.g., written symbols)."
>
> "Signal-dependent branch metric function" means "a 'branch metric function' that accounts for the signal-dependent structure of the media noise."

RDCTC at 11-12.

Marvell's constructions of "signal-dependent noise" and "signal-dependent branch metric function" is:

> "Signal-dependent noise" means "noise that is dependent on the signal."
>
> "Signal-dependent branch metric function" means "a 'branch metric function' that accounts for 'signal-dependent noise.'"

RDCTC at 11-12.

The entire term “signal-dependent noise” does not appear in the specification. However, the terms “signal-dependent” and “signal-dependence” do appear in the specification and are used as follows[10]:

> The nonstationarity of the **media noise** results from its **signal dependent** nature. Combating **media noise** and its **signal dependence** has thus far been confined to modifying the Euclidian branch metric to account for these effects. Zeng, et al., "Modified Viterbi Algorithm for Jitter-Dominated 1-D^2 Channel," IEEE Trans. Magn., Vol. MAG-28, pp. 2895-97, Sept. 1992, and Lee et al., "Performance Analysis of the Modified [M]aximum Likelihood Sequence Detector in the Presence of Data-Dependent Noise," Proceedings 26th Asilomar Conference, pp. 961-64, Oct. 1992 have derived a branch metric computation method for combating the **signal-dependent** character of **media noise**.
>
> The trellis/tree branch metric computation of the present invention is correlation-sensitive, being both **signal-dependent** and sensitive to correlations between noise samples.
>
> These covariance matrices are different for each branch of the tree/trellis due to the **signal dependent** structure of the **media noise**.
>
> Due to the **signal dependent** nature of **media noise** in magnetic recording, the functional form of joint conditional pdf $f(r_1, . . . , r_N| a_1,. . . , a_N)$ in (1) is different for different symbol sequences $a_1,. . . , a_N$.
>
> [The branch metric] is also dependent on the postulated sequence of written symbols a_i -K_1,. . . , a_i +L+K_t, which ensures the **signal-dependence** of the detector.
>
> The noise is now considered to be both correlated and **signal-dependent**.

---

[10] The term “signal-dependent” appears in both patents with and without a hyphen. No distinction is made based on the hyphen and for the sake of consistency, except where being quoted in the patents, the Court will use the term with the hyphen.

> First, $\underline{w}_i$ and σ_i^2 can be obtained directly from Equations (20) and (16), respectively, once an estimate of the **signal-dependent** covariance matrix C_i is available.
>
> It is important to point out that, due to the **signal-dependent** character of the **media noise**, there will be a different covariance matrix to track for each branch in the tree-trellis of the Viterebi-like detector.
>
> The reason for this is that the PR4 shaping filter averages noise samples from different symbols, which masks the **signal dependent** nature of the **media noise**.

'839 Patent col. 1 ln. 39-51; col. 2 ln. 9-12; col. 2 ln. 18-20; col. 4 ln. 24-27; col. 5 ln. 49-52; col. 6 ln. 38-39; col. 8 ln. 11-14; col. 10 ln. 18-21; col. 12 ln. 51-54. (emphases added). It is apparent from these sections of the specification that the only conjunction of the terms "noise" and "signal-dependent" is with the term "media noise."

Another intrinsic use of the terms is found in two of the articles cited in the specification by the patentee as "[c]ombating media noise and its signal dependent nature," which describe what is defined in the specification as the "variance dependent branch metric." In Zeng, et al., "Modified Viterbi Algorithm for Jitter-Dominated 1-D^2 Channel," IEEE Trans. Magn., Vol. MAG-28, pp. 2895-97, Sept. 1992, the abstract states that "[o]ne way to improve data capacity in magnetic recording is to increase linear density by storing magnetic transitions more closely in each track. . . . Transition noise cannot be modeled as additive noise since it is data-dependent."[11] (Docket No. 82-10 at 2). In the Zeng article, the metric that is concluded to account for the data-dependent transition noise is "ln $\sigma_k^2 + (Z_k - y_k)^2 / \sigma_k^2$ rather than $(Z_k - y_k)^2$, the standard error metric for the [Viterbi algorithm]." (Docket No. 82-10 at 3). The PHOSITA

[11] Marvell does not dispute that data-dependent and signal-dependent have the same meaning. (Docket No. 119 at 36:2-10).

would recognize this metric to be the same as what is represented in the specification as the variance dependent branch metric, equation (10).

$$M_i = \log\sigma_i^2 + \frac{N_i^2}{\sigma_i^2} = \log\sigma_i^2 + \frac{(r_i - m_i)^2}{\sigma_i^2} \tag{10}$$

'839 Patent col. 6. ln. 31-35.

Similarly, in Lee et al., "Performance Analysis of the Modified Maximum Likelihood Sequence Detector in the Presence of Data-Dependent Noise," Proceedings 26th Asilomar Conference, pp. 961-64, Oct. 1992, the abstract states "[a]s recording densities grow in magnetic storage, transition-dependent noise becomes more significant . . . . A first conclusion is perhaps to derive a new error metric that considers data-dependent noise . . ." (Docket No. 82-9 at 2). The Lee article explains what the authors mean by data-dependent media noise as follows:

> There are two main sources of data-dependent media noise. The first is non-deterministic transition shift. As the boundary of transition, inter-reaction of the magnetic material causes transition shift, **depending on write patterns**. The second is pulse amplitude fluctuation, caused by fluctuation of transition width with data pattern.

(Docket No. 82-9 at 2)(emphasis added). The Lee article then provides an error metric for data-dependent noise that is the same as what is found in equation (10) in the specification.

$$\sum_{k=1}^{M} (\log \sigma_k^2 + \frac{n_k^2}{\sigma_k^2})$$

where $n_k = z_k - y_k$

(Docket No. 89-2 at 3).

Based upon the above articles, the PHOSITA would understand the variance dependent branch metric to be a signal-dependent branch metric. The specification explains for the variance dependent branch metric the "variance depends on the written sequence of symbols." '839 Patent col. 6 ln. 17-18. From this, the PHOSITA would conclude that the noise attributable to the written sequence of symbols on the disk is "media noise", and would further conclude that the variance dependent or signal-dependent metric function is specifically stated to account for the noise from the written sequence of symbols, that is to say "media noise." This is further confirmed by the language discussing the branch metric function from which the euclidean, variance dependent and correlation-sensitive branch metric functions are derived. "[The branch metric] is also dependent on the postulated sequence of written symbols a_i -K_1,. . . , a_i +L+K_t, which ensures the signal-dependence of the detector." '839 Patent col. 5 ln. 49-52. Thus, to ensure signal-dependance, the branch metric accounts for the sequence of written symbols.

Marvell cites to the last paragraph of the specification as evidence that the patentee intended the scope of patent to be beyond magnetic media.

> While the present invention has been described in conjunction with preferred embodiments thereof, many modifications and variations will be apparent to those of ordinary skill in the art. For example, the present invention may be used to detect a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel. The foregoing description and the following claims are intended to cover all such modifications and variations.

'839 Patent col. 13 ln. 51-59. The Court is cognizant that, as stated above, "there is sometimes a fine line between reading a claim in light of the specification, and reading a limitation into the claim from the specification." *Phillips*, 415 F.3d at 1323 (citing *Comark,* 156 F.3d at1186). However, based upon the consistent pairing of "signal-dependent" with "media noise" as well as

the specification's explicit statements that metrics are signal-dependent because they account for the written sequence of symbols, the PHOSITA would conclude that "signal-dependent noise" is media noise attributable to the written sequence of symbols.

Thus, the construction of the terms "signal-dependent noise" and "signal-dependent branch metric function" is as follows:

Signal-dependent noise means "media noise in the readback signal whose noise structure is attributable to a specific sequence of symbols (e.g., written symbols)."

Signal-dependent branch metric function means "a 'branch metric function' that accounts for the signal-dependent structure of the media noise."

E. Viterbi-like

Before briefing the disputed terms, the parties had an agreed upon construction for the term "Viterbi-like." The prior agreed construction was "Viterbi-like means similar to and including the Viterbi algorithm" where the disputed term was "Viterbi algorithm." (Docket No. 74-1 at 4). Following the initial claim construction briefing and discussion at the hearing, it became apparent that the parties disputed the scope of the term "Viterbi-like."

"When the parties raise an actual dispute regarding the proper scope of these claims, the court, not the jury must resolve that dispute." *O2 Micro International Limited v. Benyon Innovation Technology Co., LTD.,* 521 F.3d 1351, 1360 (Fed. Cir. 2008)(citing *Markman*, 52 U.S. at 979). Marvell stated at the hearing that based on a construction of "Viterbi", along with the ordinary meaning of the word "like", the jury would be asked if the accused device(s) and/or method(s) contained an element that was "Viterbi-like." However, that the parties dispute, for instance, that the Fitzpatrick Patent is "Viterbi-like", makes clear there is a question of claim

scope that cannot be put before the jury. "A determination that a claim term 'needs no construction' or has the 'plain and ordinary meaning' may be inadequate when the term has more than one 'ordinary' meaning or reliance on a term's 'ordinary' meaning does not resolve the parties' dispute." *O2 Micro*, 521 F.3d at 1361. Furthermore, the question (of claim scope) as to when an algorithm is like the "Viterbi algorithm" is a question that must be answered by asking how the PHOSITA would read the claims, thus requiring the court to construe the claim term.[12] Only then, with that question answered, could a jury answer the question of whether the accused device(s) and/or method(s) contain elements that the PHOSITA would understand to be "Viterbi-like."

The term "Viterbi-like" appears only in the '839 Patent in claims 1, 4, 11, 16, 19 and 23. Claims 1 and 23 of the '839 Patent are representative of how "Viterbi-like" is used and they read as follows:

> **1**. A method of determining branch metric values for branches of a trellis for a Virterbi-like [sic] detector, comprising:
> selecting a branch metric function for each of the branches at a certain time index; and
> applying each of said selected functions to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch metric function was selected, wherein each sample corresponds to a different sampling time instant.
>
> **23**. A system for recording information on a magnetic medium, comprising:
> a write signal processing circuit for processing a plurality of data from a data source;
> a write control circuit;

[12] The term "Viterbi-like detector" can be found in Zeng, et al., Modified Viterbi Algorithm for Jitter-Dominated 1-D^2 Channel, (Docket No. 82-10 at 3), thus indicating that the term has meaning for those in the art. However, the Zeng article does not use the term in a manner such that it is helpful to this Court's construction of the disputed term.

a write head responsive to said write control circuit for receiving a plurality of signals from said write signal processing circuit, said write head for writing said signals to the recording medium;
a read control circuit;
a read head for reading said signals from the recording medium, said read head responsive to said read control circuit; and
a detector circuit for detecting a plurality of data from said read signals, said detector comprising:
a Viterbi-like detector circuit, said Viterbi-like detector circuit for producing a plurality of delayed decisions and a plurality of delayed signal samples from a plurality of signal samples;
a noise statistics tracker circuit responsive to said Viterbi-like detector circuit for updating a plurality of noise covariance matrices in response to said delayed decisions and said delayed signal samples; and
a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

'839 Patent col. 13 ln. 61 - col.14 ln. 2; col. 16 ln. 22-51.

CMU's construction of "Viterbi algorithm" and "Viterbi-like" are as follows:

> "Viterbi algorithm" means "an iterative algorithm that uses a trellis to determine the best sequence of hidden states (in this case, written symbols) based on observed events (in this case, observed readings that represent the written symbols), where the determined sequence is indicated by the best path through the trellis and is determined using branch metric values calculated for branches of the trellis."
>
> "Viterbi-like [algorithm]" means "an algorithm that is or is similar to an iterative algorithm that uses a trellis to determine the best sequence of hidden states (in this case, written symbols) based on observed events (in this case, observed readings that represent the written symbols), where the determined sequence is indicated by the best path through the trellis and is determined using branch metric values calculated for branches of the trellis."

RDCTC at 13-14.

Marvell's constructions of "Viterbi algorithm" and "Viterbi-like" are as follows:

> "Viterbi Algorithm" means "an algorithm that uses a trellis to perform sequence detection by calculating branch metrics for each branch of the trellis, comparing the accumulated branch metrics for extensions of retained paths leading to each node of the trellis at a given time, and for each node, retaining only the path having the best accumulated metric."
>
> "Viterbi-like" means "similar to and including the 'Viterbi algorithm.'" The Viterbi algorithm includes the step of calculating branch metrics for each branch of the trellis. Therefore, a "Viterbi-like" algorithm must either calculate branch metrics for each branch of the trellis or it must perform a step similar to calculating branch metrics for each branch of the trellis (in addition to other steps similar to or identical to the other Viterbi algorithm steps) such that the overall sequence detection process is similar to the Viterbi algorithm. Under Defendant's construction, a process that only calculated branch metrics for a fraction of the branches in a trellis or only compares a few paths would not be Viterbi-like.

RDCTC at 13-14.

At the hearing, the parties presented two discreet disputes as to the scope of "Viterbi" and "Viterbi-like." Those issues were whether the "Viterbi algorithm" had to calculate a branch metric for every branch of the trellis and whether the "Viterbi-like" algorithm covered the use of a post-processor as used in U.S. Patent 5,689,532 ('532 Patent or "Fitzpatrick Patent"). In the post hearing briefing and reply, Marvell indicated that "Viterbi-like" "does not require calculating branch metrics for every branch of a trellis, and it does not require a specific add-compare-select process at every node of the trellis." (Docket No 138 at 6). Rather, its construction allows for "a step similar to calculating branch metrics for each branch of the trellis" to be used for "Viterbi-like." RDCTC at 13-14. Furthermore, Marvell does not dispute that the reduced state RAM-RSE algorithm referenced in the '839 Patent at col. 1 ln. 33, col. 7 ln. 9-10, is a "Viterbi-like" algorithm. *Id*. Marvell continues to dispute, however, that the post-processor error correction circuitry in Fitzpatrick is encompassed by "Viterbi-like."

"Viterbi-like" appears throughout the specification and examples of "Viterbi-like" detectors are given in the background of the invention section.

> Traditional peak detectors (PD), such as those described in Nakagawa et al., "A Study of Detection Methods of NRZ Recording", IEEE Trans. Magn., vol. 16, pp. 1041-110, Jan. 1980, have been replaced by Viterbi-like detectors in the form of partial response maximum likelihood (PRML) schemes or hybrids between tree/trellis detectors and decision feedback equalizers (DFE), such as FDTS/DF, MDFE and RAM-RSE.

'839 Patent col. 1 ln. 26-33. Furthermore, in the detailed description section, it is stated that "[i]n the derivations of the branch metrics (8), (10) and (13), no assumptions were made on the exact Viterbi-type architecture, that is, the metrics can be applied to any Viterbi-type algorithm such as PRML, FDTS/DF, RAM-RSE, or, MDFE." '839 Patent col. 7 ln. 5-9.

During the prosecution of the '839 Patent, the examiner issued a rejection finding that the Fitzpatrick patent anticipated claims of the '839 Patent application. (Docket No. 128-2 at 22-24). Pertinent to this discussion is that the examiner found that "Fitzpatrick discloses a method for determining branch metric values of [a] trellis for a Viterbi-like detector." (Docket No. 128-2 at 23). The patentee successfully traversed the examiner's rejection with the following statement:

> Applicants have herein amended claims 1, 4, 27, and 28 to clarify that each of said selected functions is applied to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch metric function was selected, wherein each sample corresponds to a different sampling time instant. Applicants submit that Fitzpatrick does not teach, among other steps, such a step. In particular, each of the branch metrics is not determined based on a plurality of signal samples.
>
> Fitzpatrick does not specify the manner in which the branch metrics are computed. However, the Viterbi detector described in Fitzpatrick is described as an EPR4 Viterbi detector. Such a Viterbi detector computes a branch metric using:
>
> $M_i\,(r_i, a_{i-3},\ldots,a_i) = [r_i - y(a_{I-3},\ldots,a_i)]^2$
>
> where r_i is a single waveform, not a plurality of time variant signal samples.

(Docket No. 83-1 at 8).

Marvell argues that the only Viterbi detector found in Fitzpatrick is the PR4 Viterbi detector and that the post-processor that follows the PR4 Viterbi detector does not fall within the scope of the term "Viterbi-like." Additionally, Marvell points out that the branch metric the patentee cited in the prosecution history was not the error-event metric calculated by the post-processor and that the patentee ignored all aspects of the post-processor in the prosecution history. Finally, Marvell points to Fig. 5 of the Fitzpatrick patent as indicating that only the PR4 detector is referred to as being a Viterbi detector and not the EPR4 detector or the post-processor[13]

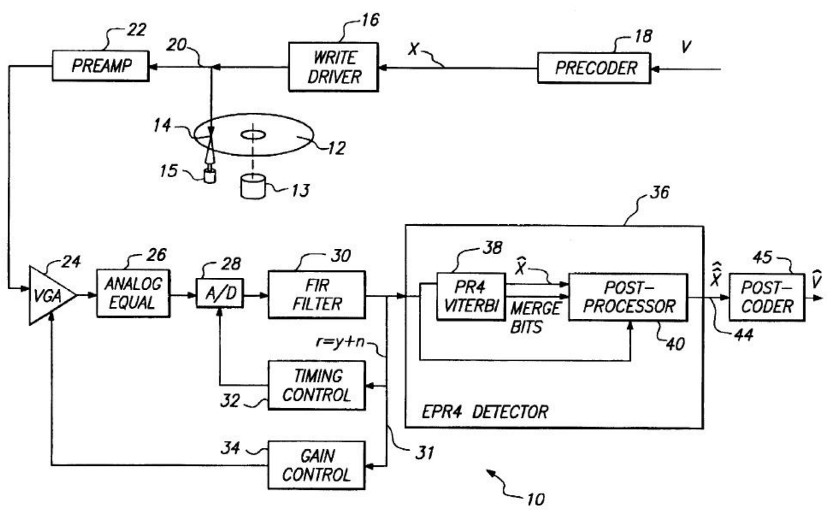


FIG. 5

'532 Patent Fig. 5. (*See* Docket No. 138 at 10-13).

---

[13]At the claim construction hearing, the Court entered into evidence Marvell Exhibit A, the 10/8/2001 email from Dr. Kavcic to Gregory Silvus. (Docket No. 106-1). The Court gives no weight to this email as it is of the type of extrinsic evidence that the PHOSITA could not be aware of since it is a personal email and it post-dates the filing and issuance of the '839 Patent. Furthermore, the email is contradicted by the intrinsic evidence, as discussed below, therefore even if the PHOSITA were aware of the contents of the email, it would be disregarded in favor of the conclusions that would be drawn from the intrinsic record.

The argument that the patentee presented to the examiner was that Fitzpatrick did not contain correlation-sensitive branch metrics because Fitzpatrick did not use a "plurality of time variant signal samples." The patentee specifically described Fitzpatrick as having an EPR4 Viterbi detector and stated that "[s]uch a Viterbi detector computes a branch metric using: M_i (r_i, a_{i-3},...,a_i) = [r_i-y(a $_{I-3}$,...,a_i)]^2" (Docket No. 83-1 at 8). This is recognized as being the Euclidian branch metric from equation (8). '839 Patent col. 6 ln. 10-14. As a result, Fitzpatrick lacked the use of a "plurality of time variant signal samples."

The specification of the Fitzpatrick patent states the invention relates to "a reduced complexity post-processor for a binary input extended partial response class 4 (EPR4) channel." '532 Patent col. 1 ln. 15-17. In discussing the prior art, it then states that "the main drawback to implementing [extending partial response class 4 with maximum-likelihood detection] within a magnetic recording system has heretofore been that the EPR4 Viterbi detector is much more complex that a PR4 Viterbi detector, and has been practically realized only at considerably greater expense." '532 Patent col. 2 ln. 8-12. The patent goes on to discuss the standard approach of implementing a Viterbi detector with a Viterbi algorithm. '532 Patent col. 2. ln. 32-55. The patent also discusses alternate approaches to implementing an EPR4 Viterbi detector and various methods leading up to the method displayed in Fig 5. '532 Patent col. 2 ln. 56 - col. 3 ln. 28. For the method using the post-processor, the patent states:

> Another implementation approach is to use a PR4 Viterbi detector, followed by a post-processor for EPR4. A post-processor for an EPR4 channel that achieves nearly maximum-likelihood performance was described by Wood, "Turbo PRML: A Compromise EPRML Detector", IEEE Trans. on Magnetics, Vol. 29, No. 6, Nov. 1993, pp. 4018-4020. In the Turbo PRML post-processor technique, PR4 equalized samples are sent to a PR4 Viterbi detector that produces a preliminary estimate of the binary input sequence. Then, the preliminary estimate is sent to the

> post-processor to produce a final improved estimate of the binary input sequence.

'532 Patent col. 3 ln 29- 40.

In the detailed description section, the patent discusses the EPR4 Viterbi detector stating:

> Typically, an EPR4 Viterbi detector is designed to *find the path through the EPR4 trellis that minimizes the squared Euclidean distance between the received samples* and the ideal EPR4 samples along the path. The output of the EPR4 Viterbi detector is a maximum likelihood sequence estimate for an EPR4 channel corrupted by independent and identically distributed Gaussian noise with zero mean.

'532 Patent col. 7 ln. 57-64. (emphasis added). In describing the post-processor in Fig. 5 the patent states:

> The post-processor **40** uses the estimated binary input sequence at the PR4 Viterbi detector output to establish a PR4 path through the EPR4 trellis. The objective of an EPR4 post-processor is to *find the path through the EPR4 trellis that minimizes the squared Euclidean distance between the EPR4 equalized samples and the noiseless EPR4 samples*, given that this path is restricted to the set of paths that deviate from the PR4 path by a sequence of non-overlapping minimum distance error-events. If the post-processor achieves this objective, then the estimated input sequence at the output of the post-processor, denoted by {x[0], x[1],..., x[j], ...}, is equal to the output of an EPR4 Viterbi detector under the conditions that only minimum distance error-events occurred and that these error-events were sufficiently far apart. The post-processor **40** produces a sequence estimate (x **44** in FIG. **5**) which is "nearly" a maximum likelihood sequence estimate for an EPR4 channel corrupted by independent and identically distributed Gaussian noise with zero mean.

'532 Patent col. 9 ln. 20-38. (emphasis added). Based on these quoted passages, the PHOSITA would conclude that the post-processor has the same objective as the EPR4 Viterbi detector and that the post-processor "nearly" produces the same result as the maximum likelihood sequence estimate for the EPR4.

Marvell cites to the following language to show that the post-processor computes too few calculations in order to be considered "Viterbi-like" :

> In the reduced-complexity post-processor described in the present invention, the post-processor calculates and compares two error-event metrics, independent of the modulation code. The post-processor uses the merge bits from the PR4 Viterbi detector to determine the "best type A" error event and the "best type B" error-event ending at a particular state. In this manner, the post-processor only considers the most-likely error-event of each type, as determined by the PR4 Viterbi detector.

'532 Patent col. 11 ln. 57-65. Although the post-processor in Fitzpatrick only calculates and compares two metrics that are first passed through a PR4 Viterbi detector, the post-processor does perform a metric calculation, comparison and selection of the most likely path through a trellis. These functions of the post-processor are what CMU pointed to in its construction of "Viterbi-like" and what the Fitzpatrick patent stated is considered to be the Viterbi algorithm. "The Viterbi algorithm is an iterative process of keeping track of the path with the smallest accumulated metric leading to each state in the trellis. The metrics of all of the paths leading into a particular state are calculated and compared. Then, the path with the smallest metric is selected as the survivor path." '532 Patent col. 7 ln. 64- col. 8 ln. 2.

After reading the specification of the Fitzpatrick Patent, the PHOSITA would conclude that an EPR4 post-processor would be understood to be "Viterbi-like" based on the examiner's rejection and the patentee's acquiescence, and Fitzpatrick's description of the post-processor. Yet, Marvell has stated that its construction would exclude the post-processor based on the limitation that "a process that only calculates branch metrics for a fraction of the branches in a trellis or only compares a few paths would not be Viterbi-like." (Docket No. 138 at 9). CMU's construction of "Viterbi-like" reflects, however, that the critical elements are that the objective

of the algorithm is to determine the best path through a trellis and that it is performed by calculating metric values. As a result, the Court will adopt CMU's construction for the term "Viterbi-like."[14] As discussed above, no construction will be needed for the term "Viterbi [algorithm]" as any definition may lead the jury to attempt to decide issues of claim construction.

As a result, it is concluded that the construction of the term "Viterbi-like" is as follows:

Viterbi-like means "an algorithm that is or is similar to an iterative algorithm that uses a trellis to determine the best sequence of hidden states (in this case, written symbols) based on observed events (in this case, observed readings that represent the written symbols), where the determined sequence is indicated by the best path through the trellis and is determined using branch metric values calculated for branches of the trellis."

IV. CONCLUSION

For the above discussed reasons, the following are adopted as the proper constructions for the disputed claim terms.

Correlation means "the degree to which two or more items (here, noise in signal samples) show a tendency to vary together."

---

[14] It is noted that '180 Patent contains language that discusses "beyond Viterbi-like detectors." '180 col. 14 ln. 10. This language, however, was only cursorily addressed in the briefing and at the hearing and the Court finds that it does not bear on whether "Viterbi-like" encompasses the post-processor found in Fitzpatrick. Furthermore, it is noted that the '180 Patent issued from a continuation-in-part application filed eleven months after the '839 Patent application and the "beyond Viterbi-like detector" language is found in the new matter added to the '180 Patent specification. Additionally, "Viterbi-like" is not found in the claims of the '180 Patent. As a result, it would be inappropriate to consider the new matter added to the specification of the '180 Patent as intrinsic evidence for construing the claims of the '839 Patent. *See Goldenberg v. Cytogen, Inc.*, 373 F.3d 1158, 1167-68 (Fed. Cir. 2004)(District court erred in considering new-matter content of a related patent as part of the intrinsic record for the patent at issue).

Correlation sensitive branch metrics means "'branch metrics' that account for 'correlation' in the signal samples by using multiple signal samples from different time instances and including at least one term in the branch metric calculation that involves multiplying signal samples from different time instances together."

A correlation sensitive metric computation update circuit means "a circuit that recalculates 'correlation-sensitive branch metrics' using statistics from the 'noise statistics tracker circuit.'"

Correlated means two items that have a tendency to vary together.

Correlated noise means "noise with 'correlation' among 'signal samples,' such as that caused by coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities."

Covariance means "the expected (mean) value of the product of (r_i-m_i) and (r_j-m_j), where r_i and r_j are observed signal samples (at time i and j, respectively) and m_i and m_j are the expected (mean) values of the samples (at time i and j, respectively)(i.e., $E[(r_i-m_i)(r_j-m_j)]$)."

Covariance matrices means "arrays of covariance of pairs of signal samples, e.g. :

$$\begin{bmatrix} \mathrm{cov}(r_i, r_i) & \mathrm{cov}(r_i, r_{i+1}) \\ \mathrm{cov}(r_{i+1}, r_i) & \mathrm{cov}(r_{i+1}, r_{i+1}) \end{bmatrix}$$

Noise covariance matrices means "covariance matrices of signal samples (where the signal samples include noise)."

Signal-dependent noise means "media noise in the readback signal whose noise structure is attributable to a specific sequence of symbols (e.g., written symbols)."

Signal-dependent branch metric function means "a 'branch metric function' that accounts for the signal-dependent structure of the media noise."

Viterbi-like means "an algorithm that is or is similar to an iterative algorithm that uses a trellis to determine the best sequence of hidden states (in this case, written symbols) based on observed events (in this case, observed readings that represent the written symbols), where the determined sequence is indicated by the best path through the trellis and is determined using branch metric values calculated for branches of the trellis."

An appropriate Order follows.

/s/ *Nora Barry Fischer*
Nora Barry Fischer
U.S. District Judge

Date: October 1, 2010
cc/ecf: All Counsel of Record